DOCUMENT RESUME

ED 355 274 TM 019 628

AUTHOR	Dolenz, Beverly
TITLE	Descriptive Discriminant Analysis: An Application.
PUB DATE	Jan 93
NOTE	28p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 28-30, 1993).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*College Students; *Discriminant Analysis; *Groups; Higher Education; High Schools; *High School Students; Predictive Measurement; Research Methodology; Social Science Research
IDENTIFIERS	Career Beliefs Inventory; *Descriptive Discriminant Analysis; *Variables; Variance (Statistical)

ABSTRACT

The use of multivariate statistics in the social sciences is essential given that the studied effects are seldom influenced by only a single variable, and instead are usually influenced by multiple variables. One multivariate technique, discriminant function analysis, has potential for many applications in the social sciences. Discriminant function analysis has two defined purposes: predictive analysis and descriptive analysis. Descriptive discriminant analysis (DDA) is described, with a discussion of methodology and interpretation using data sets and graphs from a current research project. The project involves 374 participants in three age levels (grade 9, grade 12, and college) and their responses to the Career Beliefs Inventory, a 25-subscale instrument. DDA involves the study of group separation or group differences. A linear combination of a subset of response variables is considered to maximize between-group variance of the linear function relative to the within-group variance. Five tables and three figures are included. (Author/SLD)

Descriptive Discriminant Analysis: An Application

Beverly Dolenz

Texas A&M University 77843-4225

Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX, January 28, 1993. The heuristic data employed in this example were provided by faculty from Baylor College of Medicine, as part of a symposium describing research and intervention projects of these faculty.

Descriptive Discriminant Analysis: An Application

ABSTRACT

The use of multivariate statistics in the social sciences is essential given that studied effects are seldom influenced by only one single variable, but are instead usually influenced by multiple variables. One multivariate technique, discriminant function analysis, has potential for many applications in the social sciences. Discriminant function analysis has two defined purposes; predictive and descriptive analysis. The present paper is an introductory focus on descriptive discriminant analysis (DDA), with a discussion of methodology and interpretation using data sets and graphs for clarification.

In the social sciences, many of the areas which attract interest are multivariate in nature. According to Thompson (1986), the reality in which social scientists are interested is usually one "in which the researcher cares about multiple outcomes, in which most outcomes have multiple causes, and in which most causes have multiple effects" (p. 9). In spite of this realization, many graduate programs in the social sciences carry statistic courses that focus on univariate statistics and culminate with analysis of variance (ANOVA). This trend cripples the social scientist's ability to answer more interesting research questions, or more importantly to answer research questions more accurately with multivariate statistics.

Statistical techniques that examine two or more dependent variables simultaneously are referred to as multivariate. The most frequently used multivariate techniques are multivariate analysis of variance (MANOVA), discriminant analysis, and canonical correlation analysis (Fish, 1988). Fish (1988) refers to two reasons why multivariate methods are often desirable and necessary. The first reason is statistical in nature and relates to the controlling of "experimentwise" Type I error rate. Given that statistical significance is primarily a function of group size (Wilkinson, 1992), researchers may be less concerned with controlling inflation of the experimentwise error rate and more attracted to multivariate methods for a second reason, the ability to reflect the reality of the data from which the researcher is working (Fish, 1988).

In this paper, the multivariate technique focused upon is

discriminant function analysis, and more specifically descriptive discriminant analysis (DDA) (Huberty & Barton, 1989). A brief background and definition of discriminant function analysis will be presented, followed by a discussion of interpretation. As a tool for illustration and explanation, printouts and graphs from a current research project are presented. The research project involves participants divided into three age levels--9th grade, 12th grade, college--and their responses on the Career Beliefs Inventory (CBI) (Krumboltz, 1991), a 25 subscale instrument.

Background Information

As presented by Fisher (1936), the initial intended use of discriminant analysis was for classification purposes to solve prediction problems. Given a sample of individuals from two or more populations, linear discriminant functions (LDF) can be derived such that a new individual can be placed, i.e., predicted, into the correct population based on those variables which combine to form the LDF (Huberty, 1975). In the mid 1960's, the function of discriminant function analysis was significantly extended from classification to include separation, discrimination, and estimation (Huberty, 1975).

In the current literature, the two main uses of discriminant analysis are described as predictive discriminant analysis (PDA) and descriptive discriminant analysis (DDA) (Huberty, 1975, 1984; Huberty & Wisenbaker, 1992). PDA is related to the original purpose of discriminant analysis as described by Fisher (1936) and involves classification and prediction of group membership. In a way similar to regression analysis, objects or individuals

are classified into well-defined populations based on results from multiple response measures. DDA, traditionally viewed as a follow-up to MANOVA (Huberty & Wisenbaker, 1992), involves the study of group separation or group differences. A linear combination of a subset of response variables is considered to maximize between-group variance of the linear function relative to the within-group variance (Huberty, 1975). Both components of discriminant analysis, DDA and PDA, involve multiple response variables and multiple groups of objects or subjects constituting a group membership variable (Huberty & Barton, 1989).

In discriminant analysis, direction of causation is determined by the research situation (Klecka, 1980). A PDA research situation involves the response variables being used to define group categories in a way analogous to multiple regression. The group variable, such as grade level or IQ, is treated as a dependent variable. However, when, the values on the discriminating variables are defined as dependent on the groups, the discriminant analysis is referred to as DDA and is an extension of MANOVA (Klecka, 1980).

The research example presented in the present paper fits the criteria for DDA. The 25-subscale multiple response format of the CBI is treated as the dependent variable. Multiple age levels--ninth grade, twelfth grade, college--represent the variable of separation. The question associated with this research situation is: "How do responses on the CBI differ based on age level?" In other words, based on group membership as represented by age, what response differences on the CBI are

present?

When to use DDA: Assumptions and Violations

The following assumptions for discriminant analysis are outlined by Klecka (1980):

1) two or more mutually exclusive groups are present;
2) there are at least two cases per group;
3) any number of discriminating variables are possible, provided that it is less than the total number of cases minus two;
4) discriminating variables are measured at the interval level;
5) no discriminating variable may be a perfect linear combination of other discriminating variables;
6) the covariance matrices for each group must be (approximately) equal, unless special formulas are used;
7) each group has been drawn from a population with a multivariate normal distribution on the discriminating variables.

According to Klecka (1980), it has been shown that discriminant analysis is a rather robust technique which can tolerate some deviation from these assumptions. In relation to the third assumption listed above, Huberty (1975) cautions that as the ratio of the number of discriminators to the number of individuals increases, "there is a tendency for the accuracy of (discrimination) to decrease if the coefficients determined on the first sample are applied to a second sample." When the group covariance matrices are not equal, the sixth assumption, the canonical discriminant functions and the classification equations

may be distorted (Klecka, 1980). Through SPSS DISCRIMINANT, the test for equality of the population covariance matrices is available using an approximate F statistic (Huberty & Wisenbaker, 1992). However, this test is also sensitive to deviations from multivariate normality. The seventh assumption, a multivariate normal distribution, is important for tests of statistical significance, but Lachenbruch (1975) has shown that discriminant analysis is not particularly sensitive to minor violations of the normality assumption if group sizes are relatively equal.

The Interpretation Process

The application of descriptive discriminant analysis (DDA) has historically been considered as a follow-up to MANOVA (Huberty, 1975). The primary run of a MANOVA program prior to a DISCRIMANT program is unnecessary, however, given that one-way MANOVA and discriminant analysis are the same thing (Huberty & Wisenbaker, 1992). Interested readers can "prove" this to themselves by running the SPSS programs listed in Appendix A.

Descriptive discriminant analysis (DDA) involves the comparison of different groups of individuals in terms of one or more measures (Huberty, 1975). Interpretation of the results from a discriminant analysis using statistical packages such as SPSS involves a process which enables the researcher to view the data results from several perspectives. Using a DISCRIMINANT program, the researcher is able to test the assumptions associated with discriminant analysis, the omnibus null hypothesis, as well as contrast effects (Huberty & Barton, 1989). The omnibus null hypothesis tested is that of equality of the

population centroids (Huberty, 1975). When the populations are significantly separated, subsequent and more detailed study of the group differences is possible. In discriminant analysis, statistics reported which are of interest include canonical correlations, eigenvalues, and Wilks' lambda, as well as standardized coefficients, structure coefficients, plots of group centroids, and a table of "typicality probability."

Discriminant analysis is a statistical technique which enables a researcher to study differences between two or more groups with respect to many variables at the same time (Klecka, 1980). The objective of the analysis is to form functions that maximize the separation of groups (between group variance) and minimize the dispersion of scores within each group (within group variance). One way of analyzing between group variables is based on linear discriminant functions, which are linear composites of measures on $\underline{p}$ random variables for individuals in $\underline{k}$ criterion groups (Huberty, 1975).

The number of discriminating functions derived in discriminant analysis equals the number of groups minus one or the number of discriminating variables minus one, whichever is fewer. The coefficients for the first function are derived so that the group means on the function are as different as possible. The coefficients for the second function are also derived to maximize the differences between the group means but under the added condition that values on the second function are not correlated with values on the first functions (Klecka, 1980). This process continues up to the number of unique functions which

can be derived, although some will be trivial and lack statistical significance. In order to determine the number of LDF's to retain, the SPSS DISCRIMINANT procedure outputs chi-squared results. Discriminant functions that are judged to not contribute to group separation are discarded.

Another way to determine the usefulness of a discriminant function is by reference to the canonical correlation coefficient, a measure of association that summarizes the degree of relationship between the groups and the discriminant function (Klecka, 1980). By squaring the canonical correlation coefficient (eta squared), the researcher is able to determine the percentage of variance accounted for in the discriminant function by the groups. In the example presented in Table 1, the first canonical correlation is 0.70, making eta squared equal to .49. The researcher concludes that a large 49% of the variance in scores on the discriminating variables is predictable from group membership information. If the canonical correlation had been low, 0.15 for example, we would know that only 2% of the variance was being accounted for by this discriminant function and might conclude that this function is not useful in describing differences between groups.

Insert Table 1 about here

The most common test for statistical significance of the discriminant functions is Wilks' lambda (Klecka, 1980). Wilks' lambda is an "inverse" measure, analogous to $1-r^2$, with a

mathematical minimum of zero and a maximum of one (Fish, 1988). The closer Wilks' lambda is to zero, the larger is the effect size and greater the group differences. When lambda equals 1.0, no group differences exist. The significance of lambda is tested by converting it into a approximation of either the chi-square or F distributions (Klecka, 1980). In Table 1, Wilks' lambda for the first function equals 0.416 and is statistically significant. Notice that Wilks' lambda for the second function is closer to one (as it always will be), thus accounting for less of the variance.

Although they cannot be interpreted directly, eigenvalues offer another source of interpretation for LDF's. The relative magnitudes of the eigenvalues can be used to describe the comparative value of each function (Klecka, 1980). Thus, the function with the largest eigenvalue is the most powerful discriminator, while the smallest eigenvalue is the weakest discriminator. In Table 1, the eigenvalues for the first function and second function are 0.96 and 0.22, respectively. Therefore, the first function is 4.4 times better at differentiating between groups than the second function.

In addition to asking an overall group separation question, discriminant analysis allows for investigation of more interesting questions which relate to the study of group contrasts (Huberty & Wisenbaker, 1992). Descriptions regarding structure and relative variable contribution can be presented for more specific group comparisons. Statistical packages such as SPSS DISCRIMINANT provide results for multivariate pairwise

contrasts. The test information reported includes F values and associated tail probabilities. Table 2 provides a SPSS printout of univariate statistics and F ratios.

Insert Table 2 about here

Because statistical significance is largely a function of sample size, the thoughtful researcher continues interpretation for meaningfulness of results past F and p values. For further interpretation of variable meaningfulness and contribution, the SPSS DISCRIMINANT output provides standardized coefficients, structure coefficients, and plots.

One means of interpreting the relative importance of a variable in relation to the discriminant function score is through standardized coefficients. The standardized coefficient gives the variable's contribution to calculating the discriminant score (Klecka, 1980). Using standardized coefficients as a basis for variable importance has drawbacks, however. In deriving standardized coefficients, the contribution of all of the variables are considered simultaneously.

However, a problem with standardized coefficients arises when variables have high intercorrelations, causing the intercorrelating variables to "compete" for weighted values. Consequently, a variable that would carry a high weight if considered alone may be "blocked" by a variable sharing the same discriminating information. Interpretation of this blocked variable's standardized coefficient would cause the erroneous

conclusion that it was not an important contributing variable. This is demonstrated in Table 3 in which the variable labeled "z25" carries a standardized weight of 0.07 on Function I, suggesting low variable importance. The structure coefficient (to be explained next) for z25 found in Table 4 is 0.36, suggesting a high variable contribution to the LDF. This discrepancy occurs because z25 is highly correlated with z14, r=0.49, a variable with the highest standardized and structure coefficients. Because of this intercorrelation, z25 had to compete with z14 for a standardized coefficient value.

Insert Table 3 and 4 about here

Unlike the standardized coefficient which considers all variable contribution to the LDF simultaneously, structure coefficients are simple bivariate correlations and therefore, are not affected by relationships with other variables (Klecka, 1980). Structure coefficients represent an "underlying structure" that examines the correlations between each outcome variable and scores on the LDF (Huberty & Wisenbaker, 1992). Like other correlation coefficients, structure coefficients range from -1.0 to 1.0, with correlations near zero representing little commonality and correlations near $\pm$1.0 representing high commonality between the variable and LDF. By noting those variables with the highest structure coefficients for a LDF, the function can be "named" (Klecka, 1980). From the structure matrix in Table 4, the researcher could name the first function

by determining those characteristics shared by variables z14, z15, z19, z1, z7, and z25. Compared to standardized coefficients, some have argued that structure coefficients are more stable and less influenced by sampling error (Darlington, Weinberg, & Walberg, 1973), but Monte Carlo results (Thompson, 1991) have not confirmed this view. However, it is clear that some emphasis must be placed on structure coefficients in the interpretation process; standardized weights should not be the sole basis for result interpretation.

An additional means of interpretation offered by discriminant analysis is by graphic methods. SPSS DISCRIMINANT plots each data case as a point with coordinates that are the case values on the LDF's. Depending on the number of functions, the cases will be represented on a rectangular coordinate system (for two functions) or a histogram (for one function). Present on the graph is the group centroid, which is an imaginary point which has coordinates that are the average of all the profile scores for each group on each variable (Van Epps, 1987). Group separation is graphically represented by points from each group being clustered in different coordinates.

Figure 1 shows a one function plot and Figure 2 shows a two function plot. In Figure 1, the group centroids are represented numerically underneath the histogram. In Figure 2, the group centroids are represented by asterisks. By drawing lines from each group centroid to the axis of each function, as represented in Figure 3, the researcher can achieve a clear visual aid to detect group differences. On Function I in Figure 3, the first

group (9th graders) are the most different from the third group (college). On Function II, groups 1 and 3 (9th grade and college) are very similar, but both different from group 2 (12th graders).

Insert Figure 1,2,3 about here

A final component of discriminant analysis which is helpful for the researcher is the ability to readily detect outliers through inspection of plots and probability tables. SPSS DISCRIMINANT produces a "typicality probability" table denoted by P(D/G), meaning "the probability of having the observed cases's distance, given membership in the stated group" (Huberty & Wisenbaker, 1992, p.184). As shown in Table 5, cases with discrepancies between the predicted group and actual group have an asterisk by them. Case number 10, for example, "looks" more like it would come from the second group than the first group where it actually "belongs." It is suggested that if there are legitimate outliers, the researcher may want to consider conducting analyses with and without them (Huberty & Wisenbaker, 1992).

Insert Table 5 about here

Conclusions

Discriminant analysis is a multivariate technique that provides an abundance of information and a unique set of steps

for interpretation. In this paper, descriptive discriminant analysis (DDA) was described in detail as well as the interpretation process. From the discussion, it is apparent that DDA provides the researcher with the statistical tools to answer various interesting questions, such as to what degree do groups differ and what variables are associated with these differences. The ability to inspect overall, as well as pairwise, separation provides the researcher with a more comprehensive statistical system from which to work.

References

Darlington, R.B., Weinberg, S.L., & Walberg, H.J. (1973). Canonical variate analysis and related techniques. Review of Educational Research, 43, 433-454.

Fish, L.J. (1988). Why multivariate methods are usually vital. Measurement and Evaluation in Counseling and Development, 21, 130-137.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7, 179-188.

Huberty, C.J. (1975). Discriminant analysis. Review of Educational Research, 45(4), 543-598.

Huberty, C.J. (1984). Issues in the use and interpretation of discriminant analysis. Psychological Bulletin, 95(1), 156-171.

Huberty, C.J. & Barton, R.M. (1989). An introduction to discriminant analysis. Measurement and Evaluation in Counseling and Development, 22, 158-168.

Huberty, C.J. & Wisenbaker, J.M. (1992). Discriminant analysis: Potential improvements in typical practice. In Advances in Social Science Methodology, 2, 169-208.

Klecka, W.R. (1980). Discriminant Analysis. Beverly Hills, CA: Sage.

Krumboltz, J.D. (1991). Manual for the Career Beliefs Inventory. Palo Alto, CA: Consulting Psychologists Press.

Lachenbruch, P.A. (1975). Discriminant Analysis. New York: Hafner.

Thompson, B. (1986, November). Two reasons why multivariate methods are usually vital. Paper presented at the annual

meeting of the Mid-South Educational Research Association, Memphis.

Thompson, B. (1991). Invariance of multivariate results. Journal of Experimental Education, 59, 367-382.

Van Epps, P.D. (1987, January). Heuristics to facilitate understanding of discriminant analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX. (ERIC Document Reproduction Service No. ED 284 871).

Wilkinson, R.L. (1992, January). Statistical significance testing: Alternatives and considerations. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Table 1

SPSS Printout: CANONICAL DISCRIMINANT FUNCTIONS

FUNCTION	EIGENVALUE	PERCENT OF VARIANCE	CUMULATIVE PERCENT	CANONICAL CORRELATION
1*	0.96354	81.05	81.05	0.7005107
2*	0.22526	18.95	100.00	0.4287737

AFTER FUNCTION	WILKS' LAMBDA	CHI-SQUARED	D.F.	SIGNIFICANCE
0	0.41565	315.170	50	0.0000
1	0.81615	72.932	24	0.0000

* MARKS THE 2 CANONICAL DISCRIMINANT FUNCTIONS REMAINING IN THE ANALYSIS.

Table 2

SPSS Printout: WILKS' LAMBDA (U-STATISTIC) AND UNIVARIATE F-RATIO

WITH 2 AND 371 DEGREES OF FREEDOM

VARIABLE	WILKS' LAMBDA	F	SIGNIFICANCE
Z1	0.87702	26.01	0.0000
Z2	0.92885	14.21	0.0000
Z3	0.97647	4.470	0.0121
Z4	0.95355	9.036	0.0001
Z5	0.96985	5.767	0.0034
Z6	0.98613	2.610	0.0749
Z7	0.86464	29.04	0.0000
Z8	0.97704	4.360	0.0134
Z9	0.94859	10.05	0.0001
Z10	0.99370	1.177	0.3094
Z11	0.91542	17.14	0.0000
Z12	0.97662	4.440	0.0124
Z13	0.97454	4.846	0.0084
Z14	0.79314	48.38	0.0000
Z15	0.85641	31.10	0.0000
Z16	0.97032	5.674	0.0037
Z17	0.98070	3.650	0.0269
Z18	0.99017	1.842	0.1600
Z19	0.84574	33.84	0.0000
Z20	0.98411	2.995	0.0513
Z21	0.99377	1.162	0.3139
Z22	0.93785	12.29	0.0000
Z23	0.96316	7.096	0.0009
Z24	0.97146	5.450	0.0046
Z25	0.88540	24.01	0.0000

Table 3

SPSS Printout: STANDARDIZED CANONICAL DISCRIMINANT FUNCTION

COEFFICIENTS

	FUNC I	FUNC II
Z1	0.36317	0.01428
Z2	-0.13700	0.36651
Z3	-0.24536	-0.13291
Z4	0.08006	-0.10265
Z5	-0.07200	0.42808
Z6	0.23943	-0.21339
Z7	0.29125	0.26807
Z8	-0.04233	-0.07421
Z9	0.15117	-0.12239
Z10	-0.16553	0.08535
Z11	0.22102	-0.47041
Z12	-0.19642	-0.02725
Z13	-0.18410	0.34680
Z14	0.37266	0.19157
Z15	0.36522	-0.27199
Z16	-0.20438	0.28356
Z17	0.01583	0.08027
Z18	0.20301	-0.05346
Z19	0.26472	0.35301
Z20	0.00167	-0.32484
Z21	-0.17300	-0.22716
Z22	0.20922	-0.11508
Z23	0.04992	-0.20547
Z24	0.04287	-0.04807
Z25	0.07485	0.27512

Table 4

SPSS Printout: STRUCTURE MATRIX

POOLED WITHIN-GROUPS CORRELATIONS BETWEEN DISCRIMINATING VARIABLES AND CANONICAL DISCRIMINANT FUNCTIONS
(VARIABLES ORDERED BY SIZE OF CORRELATION WITHIN FUNCTION)

	FUNC I	FUNC II
Z14	0.49936*	0.30191
Z15	0.41680*	-0.03556
Z19	0.39619*	0.37193
Z1	0.38111*	-0.03531
Z7	0.36701*	0.34474
Z25	0.36328*	0.10048
Z22	0.26144*	-0.04251
Z9	0.23148*	-0.10669
Z4	0.21420*	-0.14142
Z23	0.19576*	-0.07686
Z24	0.16741*	-0.10265
Z12	-0.15637*	0.04082
Z8	-0.15337*	-0.06094
Z3	-0.14789*	0.11581
Z20	0.12504*	-0.06920
Z2	-0.22413	0.35382*
Z11	0.27006	-0.31341*
Z13	-0.09782	0.27393*
Z5	0.12309	0.27058*
Z17	-0.07878	0.24659*
Z16	0.13370	0.24359*
Z6	0.08337	-0.18091*
Z10	0.05268	-0.12764*
Z18	0.08083	0.12702*
Z21	0.06818	-0.08906*

TABLE 5

SPSS PRINTOUT: TYPICALITY PROBABILITY

CASE NUMBER	ACTUAL GROUP		HIGHEST PROBABILITY GRP	P(D/G)	P(G/D)	2ND HIGHEST GRP	P(G/D)	DISCRIMINANT SCORES...	
1	1		1	0.5296	0.4604	3	0.3073	0.0740	-0.8046
2	1		1	0.3839	0.9536	2	0.0416	-2.0320	-1.0276
3	1		1	0.9693	0.7702	2	0.2019	-1.1466	-0.2078
4	1	**	2	0.8567	0.6992	3	0.1980	0.6123	1.1311
5	1		1	0.6835	0.7926	2	0.1986	-1.6336	0.2358
6	1		1	0.8730	0.6063	2	0.2699	-0.4167	-0.4332
7	1		1	0.8534	0.8555	2	0.1227	-1.3155	-0.6123
8	1		1	0.3792	0.9145	2	0.0832	-2.2824	-0.0790
9	1		1	0.2387	0.9592	2	0.0395	-2.5603	-0.5558
10	1	**	2	0.3203	0.8768	1	0.1023	-0.2976	2.2588
11	1	**	3	0.7535	0.5262	2	0.2797	0.6905	-0.4228
12	1		1	0.8307	0.6547	2	0.2211	-0.4579	-0.6547
13	1		1	0.5111	0.6981	3	0.1684	-0.3819	-1.2710
14	1		1	0.2874	0.9601	2	0.0379	-2.3827	-0.7731
15	1	**	2	0.6481	0.7219	1	0.2456	-0.5197	1.4632
.									
.									
.									
351	2		2	0.5105	0.8537	3	0.0857	0.5142	1.9202
352	2		2	0.7752	0.7947	1	0.1323	0.0940	1.5296
353	2	**	1	0.2166	0.7434	2	0.2542	-2.1206	1.0174
354	2	**	1	0.4527	0.9442	2	0.0481	-1.8257	-1.0856
355	2	**	3	0.5380	0.3774	1	0.3203	0.3321	-0.4667
356	2		2	0.5603	0.5594	3	0.3911	1.2106	1.0145
357	2		2	0.7262	0.6925	1	0.2682	-0.4836	1.3038
358	2		2	0.6627	0.7426	3	0.1936	0.8192	1.4334
359	2	**	1	0.4620	0.8027	2	0.1928	-1.9168	0.4782
360	2	**	3	0.9106	0.6273	2	0.2944	1.1727	0.0065
361	2		2	0.9365	0.6256	3	0.2296	0.5139	0.7969
362	2		2	0.2516	0.8455	3	0.1298	1.0876	2.1914
363	2		2	0.5748	0.6383	3	0.3121	1.1224	1.2261
364	2		2	0.4336	0.7261	1	0.2564	-0.7929	1.7004
365	2		2	0.8876	0.7419	3	0.1353	0.3782	1.2514
366	2		2	0.0881	0.9075	1	0.0860	-0.6950	2.8532
367	2		2	0.4068	0.8674	1	0.1043	-0.1846	2.1167
368	2		2	0.2019	0.9095	1	0.0743	-0.2595	2.5592
369	2		2	0.8921	0.7277	3	0.1532	0.4420	1.1983
370	2		2	0.2936	0.9079	1	0.0464	0.3777	2.3677
371	2		2	0.7486	0.7882	1	0.1547	-0.0769	1.5441
372	2		2	0.8815	0.5089	1	0.3240	0.0131	0.3358
373	2		2	0.3634	0.8923	1	0.0731	0.0597	2.2383
374	2		2	0.0060	0.9826	3	0.0100	0.5887	3.9855

** denotes mismatched case

FIGURE 1. SPSS PRINTOUT: ALL-GROUPS STACKED HISTOGRAM

```
               CANONICAL DISCRIMINANT FUNCTION I
  32+
  I
  I
F I
R 24
E I                                       2
Q I                                       2
U I                                       2   2
E 16                                      2 2 2
N I                          22  2 2   2222 2 2
C I                          22 22222222222 222
Y I                        2 22 22222222222 222 2
  8              2  12     22222212222222222222222  2
  I              1 112  2 2111122122122222122222222 2 22
  I            1 1 111 22 1111122121112111122222222 2 222   2
  I  1  11 11 1111111111111111112111112111111112211212212 222 2
   X---+---------+---------+---------+---------+---------+-------
 OUT-3.0      -2.0      -1.0        .0       1.0       2.0
   11111111111111111111111111111111222222222222222222222222222222
CLASS CENTROIDS               1          2
```

CANONICAL DISCRIMINANT FUNCTIONS EVALUATED AT GROUP MEANS (GROUP CENTROIDS)

GROUP	FUNC I
1	-0.77932
2	0.34603

SYMBOLS USED IN PLOTS:

SYMBOL	GROUP	LABEL
1	1	MALE
2	2	FEMALE

FIGURE 2. SPSS PRINTOUT: TWO-FUNCTION PLOT WITH GROUP CENTROIDS
ALL-GROUPS SCATTERPLOT - * INDICATES A GROUP CENTROID

```
                      CANONICAL DISCRIMINANT FUNCTION I

            -4.0      -2.0       .0       2.0       4.0
      X---------+---------+---------+---------+---------+-------
  OUT X
C     I
A     I
N     I
O     I
N     I
I 4.0+                                 2
C     I                            1
A     I
L     I                          2
      I                     1  1   2     2
D     I                      2     12 2  2  3
I 2.0+                          1 22 1 22
S     I                       2122122    32
C     I                  22  231 222322 2 22 2    3
R     I                  2 1  1   12 2 1222232
I     I                 1 11  1  112 *122 3232
M     I                1112 1 2 212 222312233 23  3  3    3
I  .0+               112 11  1112 23222 232332 3 33
N     I             11   1111 1 *3322221333*  33 33
A     I                11 11111111111 1123333 3 23 3 3
N     I                  112111311311112223   3333 3  3
T     I                  111 12 11121    3 3     3
      I              1        1  1 1 3333 3 33             3
F -2.0                     1          3     31  3
U     I                          1
N     I                  1            3
C     I
T     I
I     I
O -4.0
N     I
      I
 II   I
      I
      I
  OUT X
      X---------+---------+---------+---------+---------+-------
              -4.0      -2.0       .0       2.0       4.0
```

(FIGURE 2 CONTINUED)

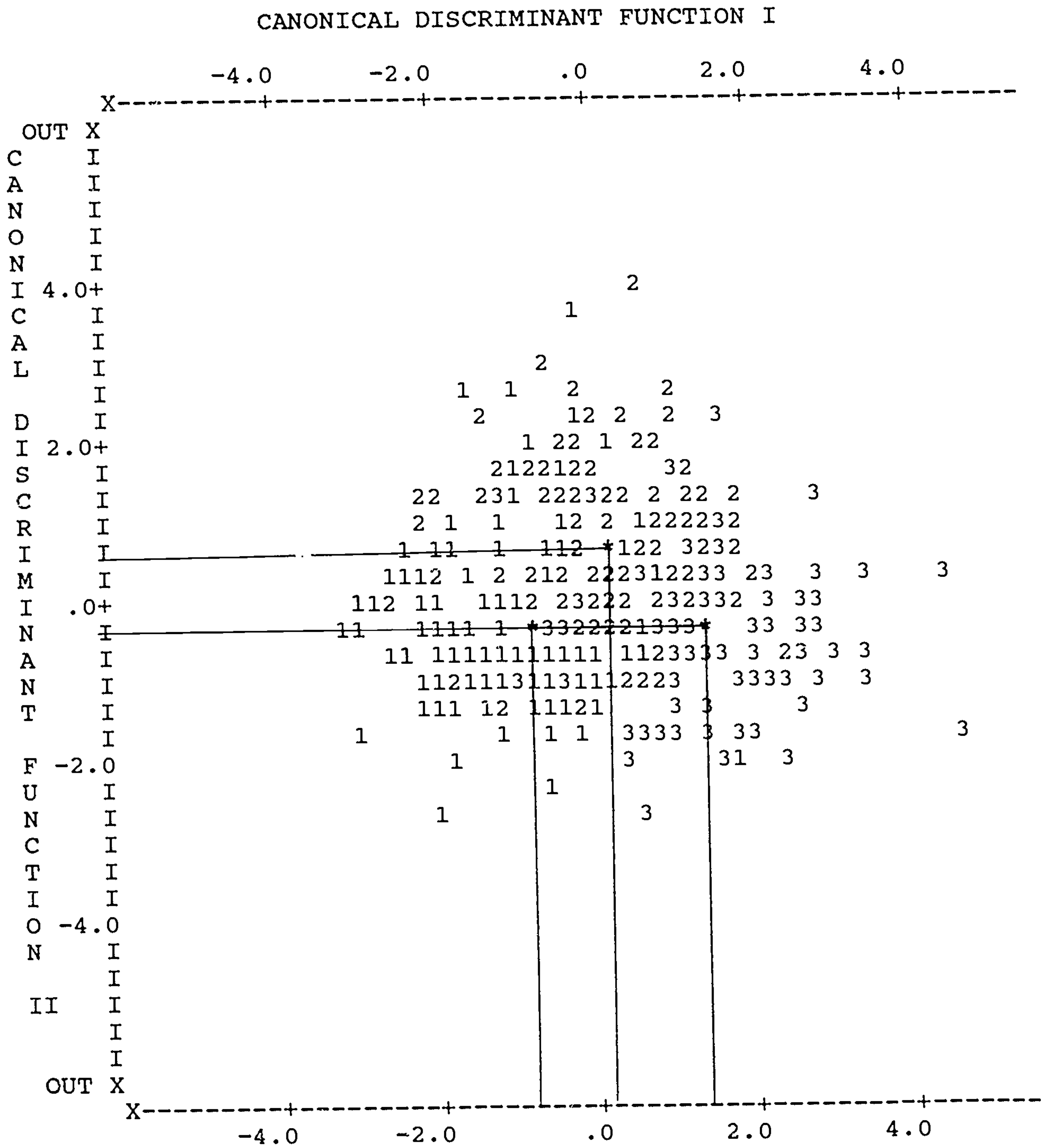
CANONICAL DISCRIMINANT FUNCTIONS EVALUATED AT GROUP MEANS (GROUP CENTROIDS)

GROUP	FUNC I	FUNC II
1	-0.89839	-0.23396
2	0.15229	0.81838
3	1.43774	-0.33559

SYMBOLS USED IN PLOTS :

SYMBOL	GROUP	LABEL
1	1	9TH GRADE
2	2	12TH GRADE
3	3	COLLEGE
*		GROUP CENTROIDS

FIGURE 3. SPSS PRINTOUT: TWO-FUNCTION PLOT WITH LINES
ALL-GROUPS SCATTERPLOT - * INDICATES A GROUP CENTROID



Appendix A

SPSS programs which "prove" that MANOVA and DISCRIMINANT are the same

```
MANOVA Z1 TO Z25 BY GRP(1,3)/
  PRINT=CELLINFO(MEANS,COV,CORR) HOMOGENEITY(BOXM)
  SIGNIF(MULTIV EIGEN DIMENR)
  DISCRIM(RAW,STAN,COR,ALPHA(.99))/ DESIGN

DISCRIMINANT GROUPS=GRP(1,3)/
  VARIABLES=Z1 TO Z25/
  PLOT=ALL/
  STATISTICS=ALL
```